DOCUMENT RESUME

ED 274 703 TM 860 580

AUTHOR Noonan, John V.; Nelson, Clare Cary
TITLE Training Instructional Designers and Subject Matter Experts in Test Development and Item Writing: Lessons from a Military Training Program.
PUB DATE Apr 86
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Criterion Referenced Tests; Instructional Materials; *Job Training; *Organizations (Groups); Specialists; Student Attitudes; *Test Construction; Test Use; *Training Methods

ABSTRACT

Although effective and efficient assessment procedures are crucial to the success of organizational training programs, the tests used for both student and program evaluation are often developed by instructional designers (IDS) and subject matter experts (SMEs) with little formal training in test development. This paper summarizes the major difficulties in training IDs and SMEs in test development and reviews strategies that facilitate this training. Training problems fall into three categories: negative attitudes about testing; mechanics of test development; and the utilization or application of tests. Negative attitudes can be diffused through increasing awareness of the benefits of testing as well as by introducing basic testing concepts and terminology. For training in test mechanics, there are excellent references such as the Psychological Corporation pamphlets. Short, activity seminars in item writing are effective methods to learn by doing (Appendix A, "Pop Quiz"). Useful references for test utilization and applications include the "Handbook for Testing in Navy Schools" and short explicit job aides, such as the "Job Aid for Embedded and Block Test Questions" (Appendix B). Involving IDs and SMEs in the decision-making process by the testing specialist is the most effective way for them to learn about test application and utilization. (BS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED274703

TRAINING INSTRUCTIONAL DESIGNERS AND SUBJECT MATTER EXPERTS IN TEST DEVELOPMENT AND ITEM WRITING: LESSONS FROM A MILITARY TRAINING PROGRAM

JOHN V. NOONAN
CLARE CAREY NELSON

FORD AEROSPACE & COMMUNICATIONS CORPORATION
HANOVER, MARYLAND

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

J.V. Noonan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1986.

BEST COPY AVAILABLE

TM 860 580

ACKNOWLEDGEMENTS

The authors would like to thank William Sweeters for permission to include his job aid for embedded and block testing in this paper (Appendix B.)

We would also like to thank Ann E. Chafin for permission to use the POP Quiz (Appendix A).

Training is fast becoming one of the most important components for organizational success. Effective and efficient assessment procedures are crucial to the success of training programs. These programs rely primarily upon criterion - referenced tests for both program and student evaluation. In many projects, staffing is such that these criterion - referenced tests are developed by Instructional Designers (IDs) and subject matter experts (SMEs) who have little formal training in test development. The quality of training programs is enhanced when designers and content specialists become facile in the field of testing. Without effective guidance and training, the IDs and SMEs develop testing and measurement skills in an unsystematic, trail-and-error fashion. Therefore, an important role for tests specialists is to become a "trainer of trainers."

The purpose of this paper is to relate experiences in training IDs and SMEs in test development. The authors will summarize major difficulties and review strategies that facilitate the training.

BACKGROUND OF PROFESSIONAL STAFF

The diversity of large groups can lead to creativity, but it also presents significant challenges. In a department of approximately one hundred experienced and talented individuals, one can expect to encounter different backgrounds and varied working styles. For the team described in this paper, another significant variable is career experience - particularly in the area of educational measurement.

The training department that was examined has a requirement for specialized personnel in the areas of instructional design and development, content expertise, measurement and evaluation , and technical application/adaptations. To meet these staffing needs, personnel are selected from government, education , and private industry. Academic backgrounds range from high school diplomas to doctorate degrees. Practical experience ranges from staff who are engaging in an initial professional position to those who are initiating second careers after military retirement. The group will include individuals accustomed to directive leadership as well as those who are comfortable in providing the direction.

While these organizational variables are not unique to a large organization, they create challenges from a test development perspective. It is critical that IDs and SMEs possess a working knowledge of educational measurement. Unfortunately, less than ten percent of the professional staff have actual test and measurement experience. The resident test specialists are faced with the critical task of both training the IDs and SMEs and promoting the measurement process.

The authors have found that staff diversity does indeed present problems in test development; however, the situation is not impossible to overcome.

In the next section, the types of problems and issues that can arise will be outlined. Then, strategies and techniques for addressing these problems will be proposed in a final section.

PROBLEMS AND ISSUES

In this section common problems and issues in teaching test development to IDs and SMEs will be discussed. The points have been separated into three groups; those related to attitudes about testing, those addressing the mechanics of test development, and those involved in the utilization or application of tests. As the points are covered, the reader will note that some of the points are general issues related to teaching anyone about test development and others more specifically apply to IDs and SMEs. The latter issues will be discussed in more detail.

The first set of issues center around attitudes toward testing. For many people, the idea of evaluating individuals by tests is offensive. Attempts to numerically represent aspects of human behavior are seen as de-humanizing. Moreover, when contrasted to measurement in the physical sciences, testing is seen as subjective and imprecise. Test development is viewed with skepticism. Unfortunately, many IDs and SMEs enter the testing endeavor with these exact preconceptions. To many who share this dislike for testing, the process is seen as a "game" that lacks substance and cannot be taken seriously. This underlying attitude can undermine attempts to develop effective evaluation.

A second source of anti-testing attitudes is the individual test anxiety of the IDs and SMEs. Perhaps because of their own negative experiences with tests, many IDs and SMEs transfer their test anxieties to the test development process. These anxieties can create a barrier that impedes training.

Finally, for many instructional designers, test development is simply not seen as very glamorous. Testing is viewed as a necessary and important component to a training program, but the rewards of training are found in curriculum design and development. Consequently, test development receives less attention and effort; tests almost become an afterthought.

A second cluster of issues includes points related to the mechanics of test development. Most of these issues are not limited to problems encountered exclusively by IDs and SMEs, but apply generally to any group that is learning test development. Leading the list are the normal problems that are expected with inexperienced item writers; such as grammatical clues to the correct answer, textbook expressions, and implausible distractors. In addition, novice writers are tempted to repeatedly use one item type or focus on lower-level cognitive tasks in order to make item writing easier. We find a good number of novice item writers who intentionally include one "throw-away" distractor on multiple-choice items. A commonly accepted rule-of-thumb for test-takers is that there is always one obviously incorrect distractor. In a peculiar realization of this prophecy, novice item writers often intentionally include a distractor that is hopelessly incorrect. In many cases, this "throwaway" is intended to be humorous -- an attempt to relieve the boredom for the examinee.

In other cases, novice item writers have a very difficult time in creating the needed number of plausible distractors.

This occurs despite the fact that in the field of instructional design, there is an available reference for distractors - the Common Error Analysis. In the Analysis phase of ISD, each objective is analyzed for common errors. Typically, subject-matter experts generate a list of errors or misunderstandings that are commonly made by trainees. The resulting Common Error Analysis is ideally suited for writing distractors. Unfortunately, many IDs and SMEs tend to ignore the Commor Error Analysis when they write test items. For many of these professionals, testing is so divorced from instructional design that they fail to realize how design efforts can assist in item writing.

As with other novice item writers, IDs and SMEs who are inexperienced at item writing tend to like hard and fast rules, rather than guidelines that require them to exercise their own judgement. A case in point is the common "rule" that one should only have one blank per completion or fill-in item. However, imagine that a CRT operator must fill-in masks on a computer screen while a client is on the telephone line, and there are fixed fields into which the operator must enter the data. For cases like these, it is often desirable to have each phone call or client contact viewed as one item with several parts. This would violate the "rule" of one blank per item, and inexperienced item writers may find this uncomfortable. Many problems arise when guidelines are given the status of fixed rules. For an even worse scenario, imagine that a set of item writing "rules" are included as formal specifications in a contract for the delivery of tests. With strict adherence to the "rules", item writers can become constrained in their attempts to develop tests

that are most appropriate for the particular content of their course.

A final issue in this cluster relates to content domains. Many measurement professionals have found content domains to be very useful in the development of criterion-referenced tests. The purpose of a content domain is to fully delineate the universe of potential items that could be written for a given objective. The domain details such things as the range of stimulus attributes, response specification, and clear rules for analyzing the correctness of all possible responses. Developing a clear and complete content domain requires a great deal of thought and effort. However, IDs and SMEs do not always recognize the need for content domain. This analysis is perceived as busywork; more often than not, the IDs and SMEs want to be given the objective and immediately initiate item writing. The danger, of course, is that there is less of a safeguard to assure that the full range of the objective is adequately represented by the particular collection of items that are written.

A final cluster of issues fall in the category of test utilization of application. The first of these is the degree of proficiency expected for the trainee. For some training programs there are extremely high expectations for the level of expertise that should be shown by the trainee. This can happen when the high degree of proficiency shown by expert job incumbents is transferred to the standard specified in the objective. A more realistic approach is to train to a novice level and not expect new trainees to display the proficiencies that can only come with experience. The consequences of not doing so can be devastating.

The program has to train to an unreasonable standard and the testing becomes so stringent that trainees have little chance of meeting criterion. Obviously, this is an issue that goes beyond testing and applies to training in general. For the purpose of this paper, it is important to point out that many SMEs struggle to understand (perhaps remember) the proper entry level of performance for new trainees.

Another difficulty results from the small sample sizes that are common in prepilots and pilots of training programs. In an ideal situation, a test developer can choose the size of the tryout and standardization sample. However, professionals in training environments usually cannot select their sample sizes. More often than not, they have to take what is given to them. The result is that they are called upon to evaluate test items when the subject pool might be as minimal as 5-10 students in a prepilot or pilot study. With such small samples it becomes very hard to interpret and present data on item statistics, reliability, and validity. Often the most sensible thing to do is to merely "eyeball" p-values. IDs and SMEs cannot really be challenged with the rich history of psychometrics and, consequently, they have difficulty understanding the scope and breadth of psychological/educational measurement. They seem interested in learning psychometric concepts, but frustration sets in when they realize that the size of the sample does not support the use of traditional psychometric measures.

A third issue in the utilization of tests relates to the setting of passing scores. We see two competing viewpoints here. The first, apparently shared by most IDs and SMEs, ties the passing score to the standard set in the objective. Thus, if the objective states that the trainee should perform some skill at 85% accuracy, then 85% is used as the passing score. If more than one objective is evaluated on the test, then one either sets a separate passing score for each objective (in effect, creating subtests that the student must pass) or sums the cut scores for each objective to arrive at a passing score for the total test. Of these two, IDs and SMEs seem to prefer the former as being instructionally purer. The second viewpoint utilizes the more traditional measurement procedures for setting passing scores, such as the Angoff (1971) or Nedelsky (1954) methods. With these procedures, proponents argue, one attains a second check on the appropriateness of the standard set in the objective. And, there is the added advantage that the actual items are examined before the passing score is set -- a "sanity check" if you will. Problems arise when these traditional procedures are used and the resulting passing score is different from the standard set in the objective. IDs and SMEs become very confused, especially if they were involved in both the task analysis that produced the objective and the item writing for the test. When IDs and SMEs are trained in traditional standard-setting procedures, a common objection is that the standard has already been detailed in the objective.

A final issue concerns the types and placement of tests. A sampling of the questions are as follows.

* Where do I place the tests? Do I put a test after each lesson in the course, or can I combine several lesson tests into a unit test?
* In what circumstances should I embed the test questions in the lessons and in what circumstances should I have a block test after a lesson or unit?
* How much content do I cover before I have a test? How long should the tests be?
* What do I include in the final course test? If a trainee has shown mastery at the lesson or unit level, why do I have to test it again at the course level? How much redundancy in testing is too much?
* How many forms of the test are needed? Why can't they take the same test twice -- this is mastery learning, isn't it?

Summarizing this section, several problems and issues that arise when attempting to train IDs and SMEs in item writing and test development have been discussed. These problems, for purposes of this presentation, have been grouped into attitudes, mechanics, and application.

STRATEGIES

The previous two sections provided an outline of the variability in experience and training among IDs and SMEs and a discussion of the major problems that arise in teaching item

writing and test development. Having identified these issues, some strategies and activities for coping with them will be discussed. As earlier, the discussion will be grouped according to the clusters of attitudes, mechanics, and application.

Regarding attitudes toward testing in general, there are several ways to diffuse the resistance. Negative attitudes are often rooted in unfamiliarity and inexperience with educational measurement. An initial solution, then, is to increase the IDs and SMEs awareness and knowledge base about testing. This is where a clear statement on the uses and benefits of testing is needed. Make no mistake, this is pure marketing. For many of these professionals, they may be having their first encounter with test development. One can disarm even those with the most anti-testing sentiments merely by presenting viewpoints in a calm and professional manner. It is important to point out advantages and to acknowledge the disadvantages of testing. The primary purpose is to create a professional image for test specialists. We have found that IDs and SMEs respond very favorably to statements about how decision-making would proceed if there were no testing -- pointing out how more subjective procedures produce unrealiable mastery-nonmastery classifications. Beyond the benefits of tests, these professionals also need to be introduced to basic concepts and terminology. Most specialists have their own technical terms and jargon and, in order to establish an initial comfort level, IDs and SMEs need to be able to understand the basic concepts and vocabulary. One should be careful

not to overwhelm the novice with advanced concepts; it is important to break the material into small, manageable units. Short seminars that deal with one concept are preferable to all-day , all - encompassing workshops. In summary, perhaps the best way to deal with negative attitudes toward testing is to strive for an open and inquisitive working environment where testing specialists invite questions and concerns. Recognize that negative attitudes may exist; do not act judgmental and and do not encourage negativism.

For training in the mechanics of item writing and test development, there are many excellent references that present clear and succint guidelines. However, IDs and SMEs do not want to be given a reading list. They know that learners learn best by doing. Nonetheless, there are several publications that are useful as reference material. The following lead our list.

1. Wesman article "Writing the Test Item" (Wesman, 1971). This is chapter from Thorndike's famous text, Educational Measurement (1971), and is a classic in the item writing literature. Wesman provides a very clear and readable summary of the major principles in item writing, including specific comments related to each of the major item forms (short answer, multiple-choice, alternate choice, matching). The chapter has dozens of good examples and nonexamples of item writing, and Wesman manages to avoid a lot of technical terms.
2. Identifying Technically Correct Test Items, Author Training Course. This is one book in the very popular

Navy Author Training Course (NAVEDTRA 10003). It is a self-paced manual with lots of good examples and practice exercises. Specific sections are devoted to true/false, multiple-choice, matching, fill-in-the-blank, short answer, and listening items.

3. The Psychological Corporation pamphlets. The Admissions and Credentialing division of The Psychological Corporation utilizes a brief, four-page pamphlet entitled, "Some Principles for Preparing Multiple-Choice Items" in training item writers for their professional certification examinations. This pamphlet provides a very straightforward and direct summary of fourteen major principles of item writing, with a "poor" and "improved" example for most of the principles. Another very popular publication from The Psychological Corporation is a Test Service Bulletin titled "A Glossary of Measurement Terms". The glossary provides clear and non-technical definitions of almost all of the major measurement terms. Simply put, it is a must for any measurement library.

4. A Technology for Test-Item Writing (Roid & Haladyna, 1982). Several chapters of this popular book are quite helpful in training item writers. Chapter 4, "Characteristics of Quality Test Items", provide the normal list of item writing guidelines and common faults in test items. However, the chapter deserves special note for its simplicity and both positive and negative examples. Chapter 5, "Measuring Skills: Rating Scales, Observations, and

Checklists", provides a very clear and readable overview of procedures for constructing rating scales, development of observations measures, and the design of behavioral checklists. Finally, Chapter 11 provides a very concise (10 pages) overview of how the Instructional Quality Inventory (IQI) (Wulfeck, Ellis, Merril, Richard, & Wood, 1978) relates to the development of quality test items. Ten pages obviously does not take the place of reading the entire IQI, but Roid and Haladayna are to be credited for their clear summary.

There are many other useful resources for item writing and test development. Most educational psychology and educational measurement textbooks address the very same issues. We only mention the ones above because we find them the most useful.

Despite these excellent references, it is important to remember that one does not become proficient at item writing by reading about how to do it: one learns by doing. Along these lines, small-group seminars can be particularly effective. In these seminars, one can select a topic that can be covered in a one to two hour session, provide an initial presentation, and, most importantly, involve the participants in active practice. An example of a possible topic for a short seminar is the "Pop Quiz" shown in Appendix A. This handout is actually a test that the participants complete. Most of the items are comprised of nonsense words, but after initial chuckles, the participants come to understand that they are to answer the items by relying

upon the subtle clues provided in the items themselves. For example, one can easily answer the following item by looking for similar root words:

4. The cresial portion of the machine is protected by the:

   a. latyz
   b. simano
   c. cresium
   d. posto

The answer is obviously "c". After the participants have taken the "Pop Quiz", each item is discussed to identify possible clues and recommend improvements. The authors have found this "Pop Quiz" to be an entertaining exercise that generates much discussion. It has a way of bringing item writing guidelines to life. Even the directions to the various item forms are intentionally written to be vague and confusing.

Another seminar topic relates to the resistance to writing content domains prior to actual item writing. In this activity, the objective is presented without a content domain and the participants are instructed to begin writing items. As the resulting items are presented and discussed in the group, the participants become surprised at the variability in how others addressed the objective. The discussion usually identifies idiosyncratic interpretations of the objective. And, the participants come to realize that the objective usually does not provide enough detail to lead the item writing.

For training on test utilization and application, there are a number of useful references that once again, we would like to cite.

1. The Handbook for Testing in Navy Schools (Ellis & Wulfeck, 1982). This is a very useful and practical guide to test development, administration, and interpretation. Again, the focus is upon the classification scheme employed in the IQI (Wulfeck, Ellis, Merril, Richard, & Wood, 1978). The Handbook provides very clear principles and examples for Remember-Level, Use-Procedure, Use-Category, Use-Rule, and Use-Principle types of objectives. They also provide a basic introduction to some simple statistics that are appropriate for a preliminary analysis of items from criterion-reference testing.

2. Hambleton's article "Advances in Criterion-Referenced Testing Technology (1982). Moving up the scale in complexity, Hambleton provides a tightly-written overview of the major issues involved in criterion-referenced testing. For purposes of training IDs and SMEs, the section on test development -- in particular his Steps in Building a Criterion-Referenced Test -- is extremely useful. He also provides a very current overview of psychometric viewpoints on reliability and validity of criterion-referenced tests, but, quite honestly, the presentation presupposes a fair degree of psychometric sophistication and is consequently too difficult for novice test developers.

In addition to the above resources, test specialists at Ford Aerospace have developed a series of short and explicit job aids. Each addresses a very specific topic; the intent is to provide

a distillation of guidelines for any specific issue. An example, provided by William Sweeters, is shown in Appendix B. This particular job aid addresses guidelines for when to use block and when to use embedded tests. Others in the series address such topics as establishing test points within a course of instruction, determining the number of items needed for an objective, and tips for developing practice items that can be used during instruction. As mentioned above, one cannot just give written materials to IDs and SMEs in order to teach them about test application and utilization -- regardless of how well-written the materials happen to be. One needs to have a test development process in which the IDs and SMEs are actively involved in addressing and solving the problems that arise. Unfortunately, and in contrast to issues related to item writing, IDs and SMEs are less likely to become directly involved in these issues if there is a testing specialist on the staff. The decisions are usually deferred to the testing specialist. However, it is best if the testing specialist does not retreat to some corner, make the decisions, and then present them to the IDs and SMEs. A more healthy strategy is to involve the IDs and SMEs in teams where they understand the rationale for the decisions. It is important to get them involved in the decision-making process, even if they initially defer to a testing specialist in a small-group session. In this way, they come to learn that there are no hard and fast rules, and the mystique of testing can gradually be overcome.

SUMMARY

In this paper, the authors have attempted to outline the major problems and issues that arise in training IDs and SMEs in item writing and test development. Special note has been made of the extensive diversity in backgrounds among these professionals. The presentation of problems and issues that arise were grouped into three clusters: those addressing attitudes toward testing, those related to the mechanics of item writing and test development, and those having to do with the application and utilization of tests. In describing some strategies for coping with these issues, the authors have tried to stress the importance of establishing and maintaining a collegial and professional working relationship among testing specialists, IDs, and SMEs. IDs and SMEs are very curious about the whole area of item writing and test development, and issues can be best addressed in an open and inquisitive working environment where testing specialists invite questions and concerns.

REFERENCES

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.). Washington, DC: American Council on Education, 508-600

Ellis, J. A., & Wulfeck, W. H. (1982). Handbooks for Testing in Navy Schools. San Diego, CA: Navy Personnel Research and Development Center

Hambleton, R. K, (1982). Advances in criterion-referenced testing technology. In Reynolds, C. R. & Gutkin, T. B. (Eds.), Handbook of school psychology. New York: Wiley

Nedelsky, L. (1954). Absolute grading standards for objective tests. Educational and Psychological Measurement, 14, 3-19.

Roid, G. H. & Haladyna, T. M. (1982). A technology for testing writing. New York: Academic Press

Wesman, A. G. (1971). Writing the test item. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.), Washington, DC: American Council on Education, 81-129.

Wulfeck, W. H., Ellis, J. A., Merrill, M. D., Richards, R. E., & Wood, N. D. (1978). The instructional quality inventory: Vol I, Introduction and overview. Sand Diego, CA: Navy Personnel Research and Development Center.

APPENDIX A

POP QUIZ

Pop Quiz

Directions: Examine each item below. If you feel the item should be answered True, please place an X on the line provided. If you feel the item should be answered False, place an O on the line provided.

I. True and False

____ 1. Only a refrix can susmit.

____ 2. A bumrule is never found in the tardy.

____ 3. A beezump is always slupper.

____ 4. A finkler is stable when the bar doe maintains its relationship constant to the frazzle.

____ 5. Samsies often eacher in order to maintain their stability.

____ 6. A dry, cool climate will affect a fubrub in such a way as to reduce its needle bop to much smaller proportions, especially when compared to its needle bop's proportions in a warm, humid climate.

Directions: Place the letter of the correct response on the line provided for each item.

II. Multiple Choise

____ 1. A frump is to the fruggle as the boz is to the: (A) biz (B) bret (C) bene (D) brollo

____ 2. To what is the fruggle related? (A) biz (B) frump (C) boz (D) brollo

____ 3. The results of the test showed that the two end products of siskum are: (A) bilge and sup (B) seep (C) smog (D) bilge

____ 4. The cresial portion of the machine is protected by the (A) latzy (B) simano (C) cresium (D) posto

____ 5. The condux has also been recognized as an: (A) bluit (B) circom (C) arlo (D) dendin

____ 6. Whenever sot and ploi became increased in proportional amounts in a sot chamber, which of the following will occur? (A) papost (B) a nimini conduct sert ax allu (C) hoi (D) kasosan

Explain why you answered each of the above as you did on the back of this page.

Evaluate each item given below. If it is a faulty item, indicate the reason why and suggestions for improvement.

1. An engine which uses explosive gas as its energy source is called an (_____) engine. (internal combustion)

2. The (_____) is obtained by dividing the (_____) by the (_____). (quotient, divident, divisor)

3. T or F Dogs are mammals, and they are the most intelligent mammals with the exception of man.

4. T or F There have been several Democratic presidents of the United States since World War I.

5. T or F Man cannot exist without oxygen.

6. T or F Women sometimes have a longer life-span than men.

7. T or F Federal aid to public schools is desirable.

8. Ethnic groups are:
   a. artifacts
   b. cultural groups
   c. racial groups
   d. external influences
   e. language groups

9. Which of the following is not a characteristic of the successful group leader?
   a. He defines the group purpose.
   b. He suggests the course of action.
   c. He shares group experiences.
   d. He opposes group interests.
   e. He justifies group actions.

10. When there is an active verb in the sentence, the subject does the acting; and when the verb is passive, the subject is acted upon. Which of the following is passive?
    a. threw
    b. hit
    c. ran
    d. was hit
    e. dropped

11. The part of speech that tells "how", "when", or "where" is an:
    a. noun
    b. pronoun
    c. adverb
    d. conjunction
    e. verb

12. Sollette and Jasquet were two explorers who
    a. sail around the world
    b. discovered the bay of Avatar
    c. becomes missionaries in order to help people
    d. has been given medal by their government

13. The power to declare war is nested in _____.

14. Sliced oranges are an excellent source of
    a. protein
    b. starch
    c. vitamin A
    d. vitamin C, if the oranges are freshly sliced

15. T or F The sum of the angles of a triangle is always 180.

16. For each of the statements to the left, choose the item on the right that goes with it. Put the letter om the line.

____ grows	a. Brazil
____ grows much coffee	b. London
____ capital of United States	c. Canda
____ capital of Great Britain	d. Washington

APPENDIX B

Job Aid for Embedded and Block Test Questions

(Note: to be used in courses that employ both block and embedded testing.)

Embedded test questions are test items which are presented to the student individually within a lesson just after their content has been learned. This is in contrast to the usual test procedure of presenting many test items as a block following a segment of instruction.

Developers should use both embedded test questions and block tests. The advantages of each are listed below:

Block Tests

- Block tests are convenient to use in traditional classroom settings.

  Block tests usually cover a fairly large segment of instruction. The mechanics of preparing the test, reproducing it, handing out papers, providing test security, correcting papers, returning tests and discussion usually takes 1-1/2 class periods. Test schedules can be distributed at the start of a course so students can plan their activities accordingly. Note that CBI changes the perspective. The mechanics of testing are handled so efficiently by a CBI system that the advantage of block test convenience does not apply to a CBI system.

- Block tests promote the integration of knowledge.

  Since content must be accessed as a whole during block tests, this promotes the integration of knowledge.
  The act of studying for a block test may aid in the building of cognitive structure.

- Block tests aid memorization

  Block tests place a load on memory and encourage students to study and organize material in a manner which facilitates its recall.

Embedded Question Tests

- Embedded questions provide for immediate reinforcement of learning.

  Since test questions are presented immediately after practice questions, the student receives the reinforcement of formally "passing" material right after it has been learned.

- Embedded questions reduce the fear of testing.

  Some students have developed a fear of testing. Since embedded questions are similar to practice questions and are presented in small doses, the fear associated with formal test sessions can often be lessened.

- Embedded questions allow instructions to flow without interruption.

  Since there are no large test breaks, embedded questions allow students to keep focused on learning new material.

Guidelines for when to use embedded test questions and when to use block tests

Add Variety

- Use both test techniques to add variety to lessons. Across lessons, try to reach a ratio of at least 1 embedded test for each 2 block tests.

- Do not ask the student too many questions. If a lesson included lots of practice questions consider using a block test to reduce the number of questions asked during the lesson.

  How much is too many questions?
  A recent study* of CAI question formats found that college students had adverse opinions about lessons in which a question appeared on every screen. A 1:1 screen to question ratio is definitely over-questioning students.

- Do not ask too few questions of the student.
  One major advantage of CAI is the capability for interactivity with the student. A simple technique to promote interactivity is to ask questions and design the system to respond according to student input. If a lesson includes mainly statements and examples without many practice items, then consider using embedded questions for testing. The same study* mentioned above found that students learned significantly more when presented with questions than when given straight text without questions.

Reduce Memory Load

- Use embedded test questions to reduce a heavy memorization load.
  If the student has to remember many detailed pieces of information, consider using embedded test questions.,

Integrate Content

- If the content to be learned must be integrated into a whole, then use block tests.

- If the content to be learned consists mainly of details to be memorized, then use embedded test items. Example: suppose the content requires learning the major components of hydraulic system (lines, pumps, valves, fluid, etc.) and specific details about each part (a valve has an input port, output port, stem, handle and a washer). In this case, the major components of different hydraulic systems might be tested as a block at the end of the lesson while the specific details about each component can be tested using embedded questions within the lesson.

Follow Hierarchy

- If the content to be learned is hierarchal, use embedded test questions to check mastery of dependent levels.

*Schloss, P. et.al. Efficacy of Four Ratios of Questions and Highlights to Text in Computer Assisted Instruction Modules. Jr. of Computer-Based Instruction, 1984, 11(4), 103-106.